

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/008,355

DATE: 12/18/2001
TIME: 16:23:46

Input Set : A:\23500440101.ST25.txt
Output Set: N:\CRF3\12182001\J008355.raw

3 <110> APPLICANT: Travis, James
4 Potempa, Jan S
5 Banbula, Agnieszka
7 <120> TITLE OF INVENTION: Dipeptidylpeptidases And Methods Of Use
9 <130> FILE REFERENCE: 235.00440101
C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/008,355
C--> 11 <141> CURRENT FILING DATE: 2001-11-08
11 <150> PRIOR APPLICATION NUMBER: US 60/246,827
12 <151> PRIOR FILING DATE: 2000-11-08
14 <160> NUMBER OF SEQ ID NOS: 26
16 <170> SOFTWARE: PatentIn version 3.0
18 <210> SEQ ID NO: 1
19 <211> LENGTH: 2139
20 <212> TYPE: DNA
21 <213> ORGANISM: Porphyromonas gingivalis
23 <400> SEQUENCE: 1

24 atgcaaatga aattaaaaag tattcttctc ggagcagccc tgctgttggg tgcttcaggg	60
26 gtagccaaag ccgacaaaagg catgtggctc ctcaacgaac tcaatcagga gaatctggat	120
28 cgaatgcgtg agctcggtt tacgctcccg ttggattcgc tctacagttt cgacaagccg	180
30 tccattgcca atgccgttgt tatcttcggt ggcggatgta cccgtatcac agtgtccgat	240
32 cagggcctga tcttaccaa ccaccactgc ggatacggtg ctatccagag ccaaagcacg	300
34 gtggatcacg actatctgcg cgatggttc gtttctcgca cgatgggtga ggagcttccg	360
36 attccgggtc tttccgtgaa gtatctgcgc aagatcgtga aggtaacggg caaggtagaa	420
38 ggacagctca agggtatcac tgacgagatg gagcgtctgc gcaaagctca ggaggtatgc	480
40 caagaactgg ccaaaaaaaga aaatgcagac gagaaccaac tctgcacatcg agagccttcc	540
42 tattccaaca acgaataactt cctcatcgtc tacgatgtat tcaaggacgt tcgtatggta	600
44 tttgctcctc ccagctctgt aggtaaagtgc ggagggcata cggacaactg gatgtggccg	660
46 cgtcacacacgg gcgacttcag cgtattccgc gtgtatgccc gtccgcacaa ccggccggcc	720
48 gaatacagca aggacaataa accctataag cccgttact tcgctccgt atccatgcaa	780
50 ggctacaagg ctgacgacta tgccatgacc atcggttcc cggggcagtac ggatcgctac	840
52 ctcacttctt ggggtgtgga agatcgtatc gaaaacgaga acaatcctcg tatcgaagtt	900
54 cgcgttatca agcaaggcat ctggaaggaa gccatgagcg cagatcagggc taccctgtatc	960
56 aaatatgcca gcaagtatgc tcagagtgc aactattggaa agaattcgat cggtatgaac	1020
58 cgcgtctcg ctcgtcttga cgtgataggt cgtaaagcgtg cccggagaaag agcattcgca	1080
60 gactggatcc gtaagaacgg caagagtgcgt gtctatggcg atgtattgtc ttctctcgaa	1140
62 aaggcttata aggaaggagc caaggccaaac cgtgagatga cttatttgag cgagacgctc	1200
64 ttccgtggta cccgggttgt tcgtttgca cagtttgcac acgcattggc tacaatcct	1260
66 gatgctcatg cccgtatcct caaatcgctt gacgacaagt acaaagacta cctccctcg	1320
68 ctcgaccgtta aggtgctgcc cccatgcgc gatattgtac gcccgcgtat ccctgcccac	1380
70 aagctcccccg atatattcaa gaatgtatc gacaagaaat tcaaaggcga cacgaagaag	1440
72 tatgcagact tcgtattcga caagagtgtg gttccattata ggcgacaaatggtt ccatgccatg	1500
74 ctcagaatcca tggacaagga aaagtttgc aaggctatcg agaaagatcc ggcagtagag	1560
76 ctttccaaga gcgtaatagc tcgtgcgc gctattcagg cccatgcgc gccaatgcc	1620
78 tatgccattg agaaggcga ggcgttttc tttgcccgtt tcgtgagat gtacccggaa	1680
80 cgtgctctgc cgagcgtatgc caacttcacc atgcgtatga gctacggctc catcaaggaa	1740
82 tatgaaccgc aggacgtgc ctggtaaacatcatacga caggcaaggg cgtattggag	1800
84 aagcaggatc ctaagagcga tgcgtttgc gtcaggaga atatcctcgac cctctccgc	1860

ENTERED

JL

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/008,355

DATE: 12/18/2001
TIME: 16:23:46

Input Set : A:\23500440101.ST25.txt
Output Set: N:\CRF3\12182001\J008355.raw

86 accaaaaact atggtcgcta tgccgagaac ggtcagctcc atatcgcttt cctatcgaac 1920
 88 aacgacatca cgggcgttaa ctccggtagc cccgtattcg ataagaacgg ccgtctgatc 1980
 90 ggtcttgcct tcgatggcaa ctggaaagct atgagtggtg acatcgagg cgaaccggat 2040
 92 ctgcagcgca caatcagcgt ggacatccgc tacgttctct tcatgattga caaatggggt 2100
 94 cagtcccccc gtctcatcca agagctgaag ttgatctaa 2139
 97 <210> SEQ ID NO: 2
 98 <211> LENGTH: 712
 99 <212> TYPE: PRT
 100 <213> ORGANISM: Porphyromonas gingivalis
 102 <400> SEQUENCE: 2
 104 Met Gln Met Lys Leu Lys Ser Ile Leu Leu Gly Ala Ala Leu Leu Leu 10
 105 1 5 10 15
 107 Gly Ala Ser Gly Val Ala Lys Ala Asp Lys Gly Met Trp Leu Leu Asn 20
 108 20 25 30
 110 Glu Leu Asn Gln Glu Asn Leu Asp Arg Met Arg Glu Leu Gly Phe Thr 35
 111 35 40 45
 113 Leu Pro Leu Asp Ser Leu Tyr Ser Phe Asp Lys Pro Ser Ile Ala Asn 50
 114 50 55 60
 116 Ala Val Val Ile Phe Gly Gly Cys Thr Gly Ile Thr Val Ser Asp 65
 117 65 70 75 80
 119 Gln Gly Leu Ile Phe Thr Asn His His Cys Gly Tyr Gly Ala Ile Gln 85
 120 85 90 95
 122 Ser Gln Ser Thr Val Asp His Asp Tyr Leu Arg Asp Gly Phe Val Ser 100
 123 100 105 110
 125 Arg Thr Met Gly Glu Glu Leu Pro Ile Pro Gly Leu Ser Val Lys Tyr 115
 126 115 120 125
 128 Leu Arg Lys Ile Val Lys Val Thr Asp Lys Val Glu Gly Gln Leu Lys 130
 129 130 135 140
 131 Gly Ile Thr Asp Glu Met Glu Arg Leu Arg Lys Ala Gln Glu Val Cys 145
 132 145 150 155 160
 134 Gln Glu Leu Ala Lys Lys Glu Asn Ala Asp Glu Asn Gln Leu Cys Ile 165
 135 165 170 175
 137 Val Glu Pro Phe Tyr Ser Asn Asn Glu Tyr Phe Leu Ile Val Tyr Asp 180
 138 180 185 190
 140 Val Phe Lys Asp Val Arg Met Val Phe Ala Pro Pro Ser Ser Val Gly 195
 141 195 200 205
 143 Lys Phe Gly Gly Asp Thr Asp Asn Trp Met Trp Pro Arg His Thr Gly 210
 144 210 215 220
 146 Asp Phe Ser Val Phe Arg Val Tyr Ala Gly Ala Asp Asn Arg Pro Ala 225
 147 225 230 235 240
 149 Glu Tyr Ser Lys [Asp Asn Lys Pro Tyr Lys] Pro Val Tyr Phe Ala Ala 245
 150 245 250 255
 152 Val Ser Met Gln Gly Tyr Lys Ala Asp Asp Tyr Ala Met Thr Ile Gly 260
 153 260 265 270
 155 Phe Pro Gly Ser Thr Asp Arg Tyr Leu Thr Ser Trp Gly Val Glu Asp 275
 156 275 280 285
 158 Arg Ile Glu Asn Glu Asn Asn Pro Arg Ile Glu Val Arg Gly Ile Lys 290
 159 290 295 300
 161 Gln Gly Ile Trp Lys Glu Ala Met Ser Ala Asp Gln Ala Thr Arg Ile

SIV 1A

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/008,355

DATE: 12/18/2001

TIME: 16:23:46

Input Set : A:\23500440101.ST25.txt

Output Set: N:\CRF3\12182001\J008355.raw

162 305 310 315 320
 164 Lys Tyr Ala Ser Lys Tyr Ala Gln Ser Ala Asn Tyr Trp Lys Asn Ser
 165 325 330 335
 167 Ile Gly Met Asn Arg Gly Leu Ala Arg Leu Asp Val Ile Gly Arg Lys
 168 340 345 350
 170 Arg Ala Glu Glu Arg Ala Phe Ala Asp Trp Ile Arg Lys Asn Gly Lys
 171 355 360 365
 173 Ser Ala Val Tyr Gly Asp Val Leu Ser Ser Leu Glu Lys Ala Tyr Lys
 174 370 375 380
 176 Glu Gly Ala Lys Ala Asn Arg Glu Met Thr Tyr Leu Ser Glu Thr Leu
 177 385 390 395 400
 179 Phe Gly Gly Thr Glu Val Val Arg Phe Ala Gln Phe Ala Asn Ala Leu
 180 405 410 415
 182 Ala Thr Asn Pro Asp Ala His Ala Gly Ile Leu Lys Ser Leu Asp Asp
 183 420 425 430
 185 Lys Tyr Lys Asp Tyr Leu Pro Ser Leu Asp Arg Lys Val Leu Pro Ala
 186 435 440 445
 188 Met Leu Asp Ile Val Arg Arg Arg Ile Pro Ala Asp Lys Leu Pro Asp
 189 450 455 460
 191 Ile Phe Lys Asn Val Ile Asp Lys Lys Phe Lys Gly Asp Thr Lys Lys
 192 465 470 475 480
 194 Tyr Ala Asp Phe Val Phe Asp Lys Ser Val Val Pro Tyr Ser Asp Lys
 195 485 490 495
 197 Phe His Ala Met Leu Lys Ser Met Asp Lys Glu Lys Phe Ala Lys Ala
 198 500 505 510
 200 Ile Glu Lys Asp Pro Ala Val Glu Leu Ser Lys Ser Val Ile Ala Ala
 201 515 520 525
 203 Ala Arg Ala Ile Gln Ala Asp Ala Met Ala Asn Ala Tyr Ala Ile Glu
 204 530 535 540
 206 Lys Gly Lys Arg Leu Phe Phe Ala Gly Leu Arg Glu Met Tyr Pro Gly
 207 545 550 555 560
 209 Arg Ala Leu Pro Ser Asp Ala Asn Phe Thr Met Arg Met Ser Tyr Gly
 210 565 570 575
 212 Ser Ile Lys Gly Tyr Glu Pro Gln Asp Gly Ala Trp Tyr Asn Tyr His
 213 580 585 590
 215 Thr Thr Gly Lys Gly Val Leu Glu Lys Gln Asp Pro Lys Ser Asp Glu
 216 595 600 605
 218 Phe Ala Val Gln Glu Asn Ile Leu Asp Leu Phe Arg Thr Lys Asn Tyr
 219 610 615 620
 221 Gly Arg Tyr Ala Glu Asn Gly Gln Leu His Ile Ala Phe Leu Ser Asn
 222 625 630 635 640
 224 Asn Asp Ile [Thr Gly Gly Asn Ser Gly Ser Pro Val] Phe Asp Lys Asn
 225 645 650 655
 227 Gly Arg Leu Ile Gly Leu Ala Phe Asp Gly Asn Trp Glu Ala Met Ser
 228 660 665 670
 230 Gly Asp Ile Glu Phe Glu Pro Asp Leu Gln Arg Thr Ile Ser Val Asp
 231 675 680 685
 233 Ile Arg Tyr Val Leu Phe Met Ile Asp Lys Trp Gly Gln Cys Pro Arg
 234 690 695 700

25/26

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/008,355

DATE: 12/18/2001
TIME: 16:23:46

Input Set : A:\23500440101.ST25.txt
Output Set: N:\CRF3\12182001\J008355.raw

236 Leu Ile Gln Glu Leu Lys Leu Ile
237 705 710
239 <210> SEQ ID NO: 3
240 <211> LENGTH: 52
241 <212> TYPE: PRT
242 <213> ORGANISM: Porphyromonas gingivalis
244 <400> SEQUENCE: 3
246 Thr Gly Gly Asn Ser Gly Ser Pro Val Phe Asp Lys Asn Gly Arg Leu
247 1 5 10 15
249 Ile Gly Leu Ala Phe Asp Gly Asn Trp Glu Ala Met Ser Gly Asp Ile
250 20 25 30
252 Glu Phe Glu Pro Asp Leu Gln Arg Thr Ile Ser Val Asp Ile Arg Tyr
253 35 40 45
255 Val Leu Phe Met
256 50
258 <210> SEQ ID NO: 4
259 <211> LENGTH: 52
260 <212> TYPE: PRT
261 <213> ORGANISM: Staphylococcus aureus
263 <400> SEQUENCE: 4
265 Thr Gly Gly Asn Ser Gly Ser Pro Val Phe Asn Glu Lys Asn Glu Val
266 1 5 10 15
268 Ile Gly Ile His Trp Gly Gly Val Pro Asn Glu Phe Asn Gly Ala Val
269 20 25 30
271 Phe Ile Asn Glu Asn Val Arg Asn Phe Leu Lys Gln Asn Ile Glu Asp
272 35 40 45
274 Ile His Phe Ala
275 50
277 <210> SEQ ID NO: 5
278 <211> LENGTH: 734
279 <212> TYPE: PRT
280 <213> ORGANISM: Shewanella putrefaciens
282 <400> SEQUENCE: 5
284 Met Ala Ser Gln Ala Leu Gly Phe Leu His Gln Asn Gly Leu Asn Thr
285 1 5 10 15
287 Met Lys Lys Trp Leu Leu Ser Val Ala Val Ala Ser Phe Ala Ser
288 20 25 30
290 His Ala Asp Glu Gly Met Trp Gln Pro His Gln Leu Pro Ala Met Ala
291 35 40 45
293 Asp Val Leu Lys Ala Lys Gly Leu Glu Ile Asp Ala Lys Ser Ile Ser
294 50 55 60
296 Lys Leu Thr Glu Phe Pro Met Asn Ala Val Ile Ser Leu Gly Gly Cys
297 65 70 75 80
299 Thr Ala Ser Phe Val Ser Pro Lys Gly Leu Val Val Thr Asn His His
300 85 90 95
302 Cys Ala Tyr Gly Ser Ile Gln Tyr Asn Ser Thr Pro Glu Lys Asn Leu
303 100 105 110
305 Leu Gln Asp Gly Phe Leu Ala Lys Thr Phe Ala Asp Glu Leu Pro Ala
306 115 120 125

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/008,355

DATE: 12/18/2001
TIME: 16:23:46

Input Set : A:\23500440101.ST25.txt
Output Set: N:\CRF3\12182001\J008355.raw

308 Ala Pro Gly Ser Arg Val Tyr Val Thr Glu Asp Val Thr Asn Val Thr
309 130 135 140
311 Glu Arg Val Lys Ala Gly Leu Glu Asn Lys Thr Gly Arg Glu Phe Tyr
312 145 150 155 160
314 Gln Gly Val Glu Asn Gln Glu Lys Ala Leu Val Ala Glu Cys Glu Lys
315 165 170 175
317 Asp Glu Gly Tyr Arg Cys Gln Val Tyr Ser Phe His Gly Gly Leu Glu
318 180 185 190
320 Tyr Tyr Leu Val Lys Gln Leu Glu Ile Arg Asp Val Arg Leu Val Tyr
321 195 200 205
323 Asn Pro Ala Gly Ser Val Gly Lys Tyr Gly Gly Asp Val Asp Asn Trp
324 210 215 220
326 Met Trp Pro Arg His Thr Gly Asp Tyr Ser Phe Tyr Arg Ala Tyr Val
327 225 230 235 240
329 Ser Lys Asn Gly Lys Pro Ala Glu Phe Ser Ala Asp Asn Val Pro Tyr
330 245 250 255
332 Glu Pro Lys Ser Phe Leu Lys Val Ser Ala Lys Gly Val Ser Glu Gly
333 260 265 270
335 Asp Phe Val Met Val Ala Gly Tyr Pro Gly Arg Thr Asn Arg Tyr Arg
336 275 280 285
338 Thr Ala Thr Glu Val Gln Asn Glu Phe Glu Trp Ala Tyr Pro Glu Gly
339 290 295 300
341 Lys Met Leu Arg Glu Arg Phe Ile Glu Ile Ile Lys Ala Thr Ala Pro
342 305 310 315 320
344 Glu Gly Ser Asp Glu Arg Ile Lys Tyr Glu Ser Gln Ile Ala Gly Leu
345 325 330 335
347 Ala Asn Tyr Ala Lys Asn Phe Thr Ser Met Ile Glu Phe Tyr Gly Lys
348 340 345 350
350 Ser Thr Met Leu Ala Asp Arg Lys Ala Leu Glu Ala Lys Leu Ala Glu
351 355 360 365
353 Trp Ile Ala Lys Asp Ser Ser Arg Glu Ala Lys Tyr Gly Lys Thr Leu
354 370 375 380
356 Ala Glu Leu Asp Ala Leu Ile Ala Lys Ser Lys Ala His Gln Glu Arg
357 385 390 395 400
359 Asp Met Ile Leu Ser Tyr Ile Ser Ser Thr Thr Met Leu Pro Thr Ala
360 405 410 415
362 Asn Asn Leu Tyr Arg Leu Ala His Glu Lys Gln Leu Pro Asp Met Gln
363 420 425 430
365 Arg Glu Pro Gly Phe Gln Asp Arg Asp Met Thr Arg Phe Lys Ala Ser
366 435 440 445
368 Met Glu Arg Ile Asp Arg Arg Tyr Ala Ala Ser Val Asp Lys Ala Val
369 450 455 460
371 Leu Phe Asp Met Leu Lys Arg Tyr Ala Ala Leu Pro Glu Ala Gln Arg
372 465 470 475 480
374 Leu Pro Ala Met Asp Lys Ala Phe Gly Ile Asp Asn Lys Val Asn Glu
375 485 490 495
377 Ala Lys Leu Ala Lys Thr Leu Asp Lys Met Tyr Ala Lys Thr Glu Leu
378 500 505 510
380 Gly Asn Lys Asp Val Arg Leu Ala Trp Met Glu Lys Ser Val Asp Asp

VERIFICATION SUMMARY
PATENT APPLICATION: US/10/008,355

DATE: 12/18/2001
TIME: 16:23:47

Input Set : A:\23500440101.ST25.txt
Output Set: N:\CRF3\12182001\J008355.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application No
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:993 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:10
L:1006 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:11
L:1019 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:12
L:1032 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:13
L:1045 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:14
L:1058 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:15
L:1071 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:16
L:1084 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:17
L:1170 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:25
L:1183 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:26